Received: 11 January 2017 | Accepted: 3 July 2017

DOI: 10.1002/hyp.11272

WILEY

RESEARCH ARTICLE

Potential of time-lapse photography for identifying saturation area dynamics on agricultural hillslopes

Rasmiaditya Silasari[1,2] | Juraj Parajka[1,2] | Camillo Ressl[3] | Peter Strauss[4] | Günter Blöschl[1,2]

[1] Centre for Water Resource Systems, Vienna University of Technology, Karlsplatz 13, Vienna 1040, Austria

[2] Institute of Hydraulic Engineering and Water Resources Management, Vienna University of Technology, Karlsplatz 13/222, Vienna 1040, Austria

[3] Department of Geodesy and Geoinformation, Vienna University of Technology, Gußhausstraße 27-29/E120, Vienna 1040, Austria

[4] Institute for Land and Water Management Research, Federal Agency for Water Management, Pollnbergstrasse 1, Petzenkirchen 3252, Austria

Correspondence
Rasmiaditya Silasari, Centre for Water Resource Systems, Vienna University of Technology, Karlsplatz 13, Vienna A-1040, Austria.
Email: silasari@waterresources.at

Funding information
Austrian Science Fund, Grant/Award Number: Vienna Doctoral Programme on Water Resource System; Austrian Science Fund (FWF), Grant/Award Number: DK W1219-N22

Abstract

Mapping saturation areas during rainfall events is important for understanding the dynamics of overland flow. In this study, we evaluate the potential of high temporal resolution time-lapse photography for mapping the dynamics of saturation areas (i.e., areas where water is visually ponding on the surface) on the hillslope scale during natural rainfall. We take 1 image per minute over a 100 × 15 m^2 depression area on an agricultural field in the Hydrological Open Air Laboratory, Austria. The images are georectified and classified by an automated procedure, using grey intensity as a threshold to identify saturation area. The optimum threshold T is obtained by comparing saturation areas from the automated analysis with the manual analysis of 149 images. T is found to be highly correlated with an image brightness characteristic defined as the greyscale image histogram mode M (Pearson correlation $r = 0.91$). We estimate T as $T = M + C$ where C is a calibration parameter assumed to be constant during each event. The automated procedure estimates the total saturation area close to the manual analysis with mean normalized root mean square error of 9% and 21% if C is calibrated for each event and taken constant for all events, respectively. The spatial patterns of saturation are estimated with a geometric mean accuracy index of 94% as compared to the manual analysis of the same photos. The patterns are tested against field observations for one date as a preliminary demonstration, which yields a root mean square error of the shortest distance between the measured boundary points and the automatically classified boundary as 23 cm. The usefulness of the patterns is illustrated by exploring run-off generation processes of an example event. Overall, the proposed classification method based on grey intensity is found to process images with highly varying brightnesses well. It is more efficient than the manual tracing for a large number of images, which allows the exploration of surface flow processes at high temporal resolution.

KEYWORDS

hillslopes, image classification, overland flow, saturation area, time-lapse photography

1 | INTRODUCTION

The mapping of saturation areas in catchments has been a topic of considerable interest since Dunne and Black (1970) performed their field studies to attribute overland flow to direct precipitation and return flow. Knowledge of the saturation areas in catchments is essential not only for understanding run-off generation processes (Kirkby, 1975; Weyman, 1973) but also for understanding their spatial patterns, which gives insight into the organizational structure of catchments (Grayson, Blöschl, Western, & McMahon, 2002; Schulz, Seppelt, Zehe, Vogel, & Attinger, 2006) and provide guidance on upscaling locally observed processes (Blöschl & Sivapalan, 1995; Grayson & Blöschl, 2001; Western, Blöschl, & Grayson, 1998).

The recent interest in hydrological connectivity in the landscape (Antoine, Javaux, & Bielders, 2009; Bracken & Croke, 2007; Bracken et al., 2013; Jencso et al., 2009) has brought about a corresponding interest in mapping the patterns of saturation. From these patterns, connectivity can be quantified by probabilistic modelling (Allard & Group, 1993; Western, Blöschl, & Grayson, 2001) that allows unravelling controls on run-off generation such as vegetation patterns, hillslope morphology, and rainfall forcing (Peñuela, Javaux, & Bielders, 2015; Reaney, Bracken, & Kirkby, 2014; Yang & Chu, 2013). Most of

these studies relied heavily on numerical modelling of saturation areas, and there is a shortage of validation to detailed real-time observation, which is difficult due to logistic constraints in the field and the dynamic nature of the saturation areas (Blazkova, Beven, & Kulasova, 2002; Meerkerk, van Wesemael, & Bellin, 2009). Therefore, procedures that can provide detailed spatial information of saturation patterns at high temporal resolution are needed.

Although Dunne and Black (1970) mapped saturation areas manually, less labour-intensive methods have in the meantime been developed, such as remote sensing techniques using airborne radar and cameras (Barrette, August, & Golet, 2000; Brun et al., 1990), Lidar (de Alwis, Easton, Dahlke, Philpot, & Steenhuis, 2007; Solomons, Mikhailova, Post, & Sharp, 2015), terrestrial photography by optical (Keys, Jones, Scott, & Chuquin, 2016; Orlandini et al., 2012), and thermal cameras (Glaser et al., 2016; Pfister, McDonnell, Hissler, & Hoffmann, 2010). Also, vegetation patterns have been mapped as a surrogate for saturation (Güntner, Seibert, & Uhlenbrook, 2004; Kulasova, Beven, Blazkova, Rezacova, & Cajthaml, 2014; Rogger et al., 2012). If one is interested in high temporal and spatial resolutions, terrestrial remote sensing methods are usually preferred.

Terrestrial remote sensing methods can be based on visual or infrared (IR) cameras. The choice of method will depend on the spatial scale and the process understanding of interest. For identifying flow sources and pathways, IR imagery can provide indications of saturated areas and any groundwater exfiltrating (Blume & van Meerveld, 2015; Deitchman & Loheide, 2009; Mundy, Gleeson, Roberts, Baraer, & McKenzie, 2017). If one is interested in the actual ponding and overland flow associated with surface saturation, visual images are more straightforward in terms of interpretation (Orlandini et al., 2012; Yang & Chu, 2013). There are many low-cost options of high-resolution optical cameras that can observe up to hundreds of square kilometres (Aschenwald, Leichter, Tasser, & Tappeiner, 2001; Salvatori et al., 2011). Thermal cameras of similar resolution are usually significantly more expensive, and they are commonly used for smaller areas of a few square metres (de Lima, Abrantes, de Lima, & de Lima, 2015; de Lima, Abrantes, Silva, de Lima, & Montenegro, 2014; Röper, Greskowiak, & Massmann, 2014) when used in a terrestrial mode.

Based on these considerations, visual terrestrial imagery is appealing for monitoring saturation areas that provides complementary information to hydrograph measurements and thus helps better understand hydrological processes. Rossi and Ares (2012) utilized digital camera imagery to capture the microtopographic effect on run-off movement and depression storage without disturbing the experimental field, a common problem that may produce artefacts in the observed flow characteristics. Similarly, Orlandini et al. (2012) used camera imagery to evaluate the dominant classes of flow direction methods in terrain-based flow direction predictions. However, all of these studies have been performed in small scale (<10 m^2) experimental settings, whereas, interestingly, hillslope scale studies of saturation areas during natural rainfall events are rare. Because overland flow processes occur over a wide range of spatial and temporal scales (Blöschl & Sivapalan, 1995), observation methods on the hillslope scale are crucial.

When observing spatial patterns of saturation area during rainfall, fine temporal sampling is important due to the relatively fast dynamics (Skøien, Blöschl, & Western, 2003). This requires time-lapse observations at a fine temporal resolution that produces a large number of images to be processed, a difficult task for manual classification. Automated procedures are faster and more objective with similar accuracy (Bauer & Strauss, 2014; Thenkabail, 2015), but they have mainly been developed for satellite imagery over areas of square kilometres (de Alwis et al., 2007; Dribault, Chokmani, & Bernier, 2012). An exception is Keys et al. (2016) who performed automated classification of an inundated river floodplain using RGB thresholds and obtained excellent agreement with model simulation results (R^2 = 0.94).

The goal of this study is to evaluate the potential of high temporal resolution time-lapse photography for mapping the dynamics of saturation areas at the hillslope scale during natural rainfall events. In this study, "saturation area" is defined as the area where water is visible on the surface, either ponding or as overland flow. This study goes beyond the previous studies by dealing with highly varying image brightness and very high temporal resolution (i.e., taking image at minute instead of hourly intervals as Keys et al., 2016). We propose an automated classification method that we test against manual classification and field measurements. We also illustrate the method's potential in capturing the spatial dynamics of saturation areas (i.e., ponding, run on, expansion, and contraction) at high temporal resolution for hydrological process studies.

2 | MATERIALS AND METHODS

2.1 | Study area

The observation site (Figure 1) is located in the Hydrological Open Air Laboratory, Petzenkirchen, Lower Austria (Blöschl et al., 2016; Exner-Kittridge et al., 2016). The Hydrological Open Air Laboratory is a 65.8 ha research catchment that is drained by a small perennial stream supplied by springs and tile drain outflow (Exner-Kittridge et al., 2016). The mean annual precipitation and temperature during 2014–2015 were 700 mm/year and 10.7 °C, respectively.

The observed depression area (from here on referred to as a thalweg, see pale red "Area of interest" in Figure 1 inset) is part of an agricultural field adjoining the stream, which was cultivated by winter wheat during the observation period presented in this paper (August 2014–July 2015). The soil type along the thalweg varies from cambisol (upstream part), kolluvisol (middle part), to planosol (downstream part). The observed saturation area in this study is approximately 100 × 15 m^2 in size with a mean slope of 4%. The field of the observed area was drained in the 1950s due to its low soil permeability (Eder, Strauss, Krueger, & Quinton, 2010) with the drain pipe located approximately 1.5 m under the surface.

From the terrain characteristics and the estimated location of the main drain pipe under the thalweg, we assume that a stream used to flow along the thalweg before the field was drained and developed into a depression area at present time. Due to the topography and the low soil permeability, overland flow occurs along the thalweg from saturation excess during the wet season and infiltration excess during heavy summer storms (Blöschl et al., 2016). Aside from local ponding, the overland flow on the thalweg comes from rainfall excess run-off of the uphill fields that flows through the valley in the middle of the

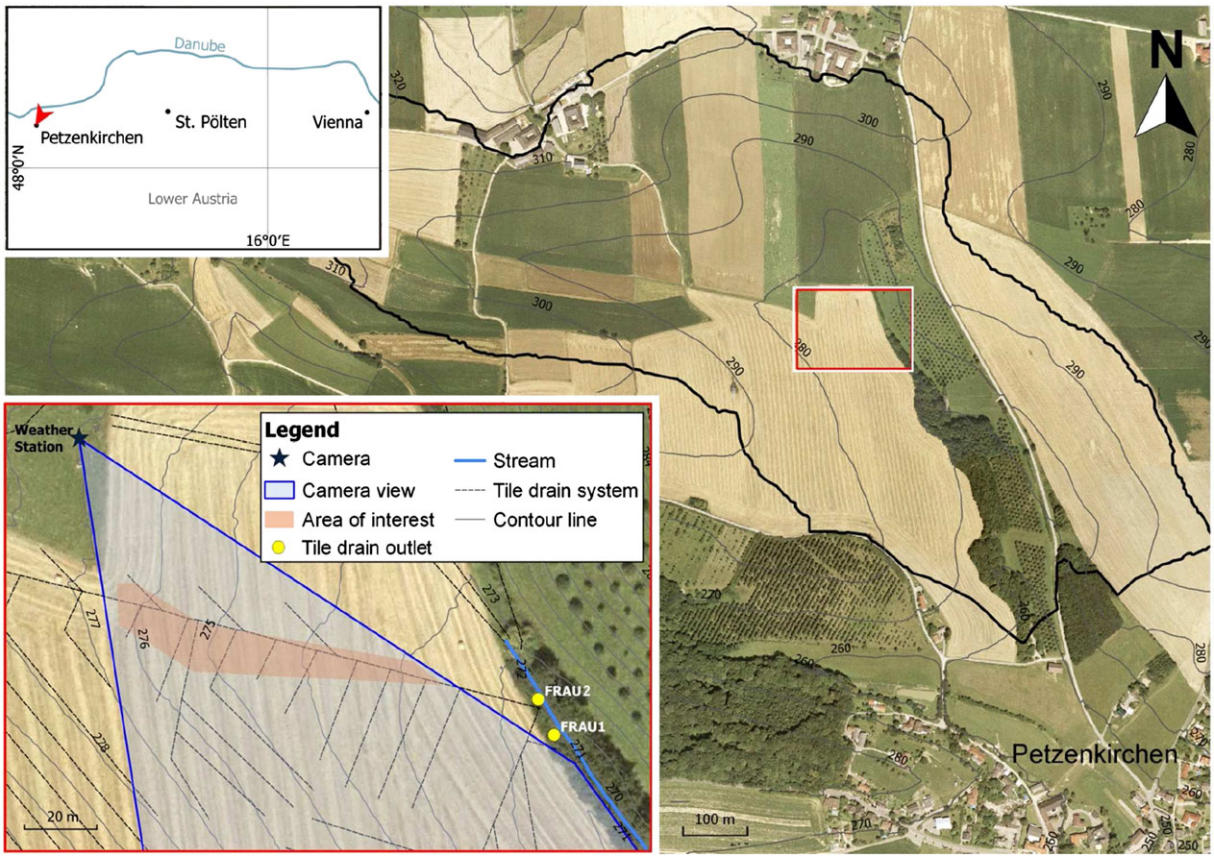

FIGURE 1 Observation site at Hydrological Open Air Laboratory, Petzenkirchen, Lower Austria. Detailed map of observed area inside the red box is shown in lower left inset

catchment (see Figure 1) to the thalweg. The contributing uphill area covers approximately 29.5 ha.

2.2 | Hydrological measurements

At a weather station next to the thalweg, rainfall is measured at 1 min resolution. Discharge of the tile drain system is monitored by two flumes (Frau1and Frau2, see Figure 1 inset) at the location where the drains enter the stream with each having different hydrograph characteristics. Frau1 outflow is not perennial (i.e., only flows during events) and shows similar response as surface run-off hydrographs (as ascertained during field work), whereas Frau2 outflow is perennial, probably originating from a perched aquifer (Exner-Kittridge et al., 2016).

In this paper, we use Frau1 outflow (mean annual outflow 0.1 l/s, 2014–2015) as an indicator of surface run-off occurrence due to the unavailability of direct surface run-off measurements during the camera observation period. From the past data of surface run-off measured at an erosion gully station near the thalweg, we observe that the flow reactions at both the erosion gully and Frau1 started almost together. This is the basis of our consideration of using Frau1 as a surrogate for surface run-off in the hydrological analysis.

2.3 | Time-lapse imagery observation

For the image observations, we chose a network camera (Sanyo 2 Megapixels VCC-HD5600, see Table 1) commonly used for surveillance due to its feature of remotely-controlled pan-tilt-zoom that provides easy control over the camera view. Based on experience in Parajka, Haas, Kirnbauer, Jansa, and Blöschl (2012), we installed the camera inside an outdoor housing unit equipped with electric fan and heating system and mounted it on the weather station mast 8 m above the ground (Figure 2). We switched the autofocus off and applied no zoom to retain similar image resolution throughout the observation period.

To obtain visual observations with high temporal resolution matching the resolution of the hydrological measurements (i.e., tile drain outflow and rainfall), we set the sampling rate to one picture per minute. The JPG image files were transferred via a fibre-optic network connection to a computer at the nearby building of the Institute for Land and Water Management Research (BAW-IKT). We monitored the area for 12 months (August 2014–July 2015) with active recording between civil dawn and civil dusk (i.e., about 30 min before sunrise and after sunset) to take sufficiently bright images.

The images presented in this paper were selected on the basis of rainfall events to find overland flow occurrence, which was especially visible in the images during low vegetation (August 2014–April 2015). The images during these selected events were analysed in more detail by image processing.

2.4 | Image processing

Figure 3 gives an overview of the steps involved in estimating the saturation areas from the photographs. The methodology comprises

TABLE 1 Camera specifications

Type	Sanyo VCC-HD5600
Resolution	2 megapixels
Image sensor	1/2.5″ progressive scan CMOS
Image compression	MJPEG/H.264
Image size	1280 (H) × 720 (V) pixels
Lens	10× optical zoom 16× digital zoom Focal length: f = 6.3–63 mm F number: F1.8–2.5
View angle (horizontal, vertical, diagonal)	Horizontal (16:9): 5.5–50.6° Vertical (16:9): 3.1–29.3° Diagonal (16:9): 6.2–57°
Minimum illumination (approx.)	Colour 2.0 lux, B/W 0.1 lux
Environmental condition (operating)	Temperature: −10 to +50 °C humidity: Within 90% RH
Power consumption	29 W (24 V AC)

two parts: (a) threshold optimization (Figure 3, left), which estimates an optimum threshold to produce automated classification results most similar to the reference (manual classification result), and (b) image processing (Figure 3, right), which classifies the saturation areas of the time-lapse images using the optimum threshold value. Because each event had different saturation area extent and natural lighting characteristics, the optimization was done separately for each event. The pixels were classified into saturated pixels (water ponding on the surface or overland flow) and into pixels where this was not the case. For brevity, these classes were termed wet and dry pixels in the image processing with no relation to degrees of soil moisture. Saturation area in the image was identified as a bright area in the middle of the field (Figure 4). It appears bright because the water film on the soil surface has higher light reflectance than its surrounding (Dobos, 2006). Katsaros, McMurdie, Lind, and DeVault (1985) suggested that, beside sunshine, cloudy skies can increase the light reflectance of the water surface, which was also the case in this study.

2.4.1 | Image georectification and cropping

Georectification transforms the oblique photograph into an orthoimage using orientation parameters of the mounted camera and a digital elevation model (DEM) as inputs (McGlone, Mikhail, Bethel, & Sensing, 2004). We estimated the orientation parameters by placing 15 ground control points (GCPs) throughout the camera field of view and measured the GCP coordinates with differential GPS (TOPCON GRS-1 with PG-A1 antenna). From the reference picture taken on February 27, 2015, we obtained the corresponding image coordinates (in pixels) of the GCPs and computed the orientation parameters using bundle block adjustment (McGlone et al., 2004). The DEM has a horizontal resolution of 0.5 m and was derived from an airborne laser scanning (ALS) campaign in March 2010.

Using the orthorectification method of McGlone et al. (2004), we produced 4400 × 3600 pixels orthoimages (5 cm pixel resolution). These were cropped to focus on the region of interest around the depression area and to remove irrelevant objects (i.e., buildings and trees), which may complicate the wet pixel classification process. The cropping resulted in a 570 × 2105 pixels orthoimage (real area 0.3 ha). A comparison between the GCP coordinates from the orthoimage and the GPS measurements gave a mean error, root mean square error (RMSE), and standard deviation of 11, 17, and 14 cm, respectively. These errors are likely caused by terrain changes between the time of the ALS campaign (2010) and the time of the reference image (2015), which are not reflected in the DEM.

2.4.2 | Manual classification

For the manual wet pixel classification of the orthoimage, we used the graphic editor software GIMP (GNU Image Manipulation Program), similar to Bauer and Strauss (2014). An initial wet pixel classification

FIGURE 2 On-site camera set-up. Left: Installation on the weather station mast. Right: Camera view

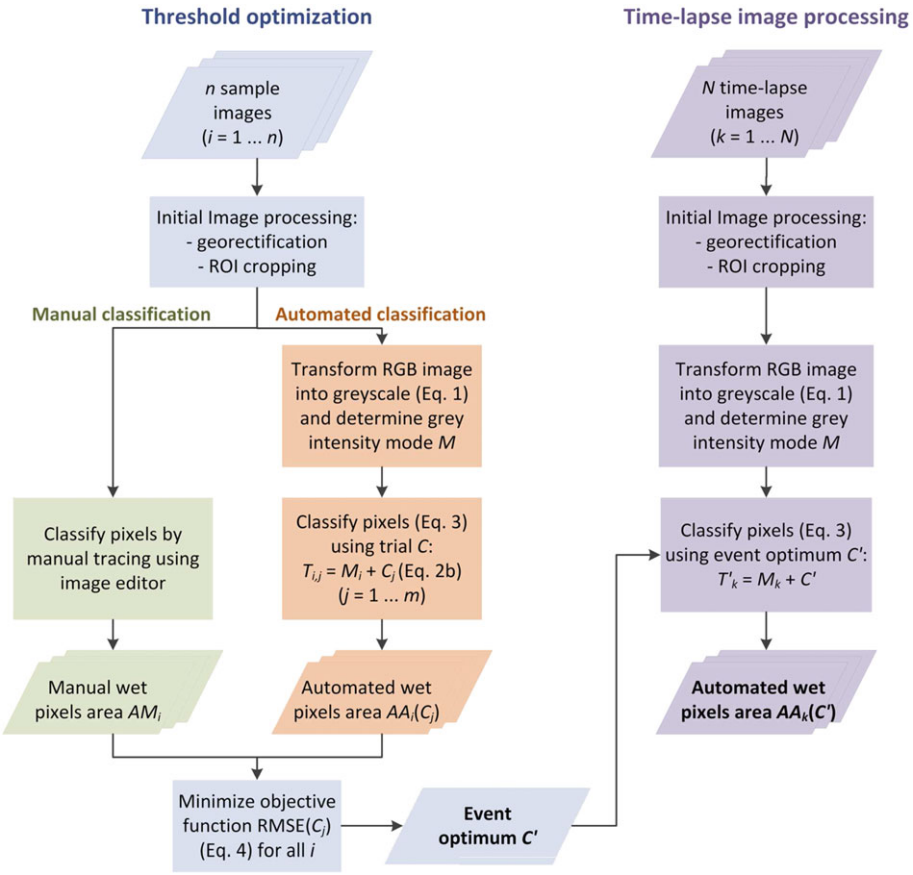

FIGURE 3 Flowchart of time-lapse image analysis for one event. Threshold optimization (left) to obtain the optimum threshold constant C' for each event is based on a subset of n sample images (those that are classified manually), whereas the final time-lapse image processing (right) is performed for the complete set of N images

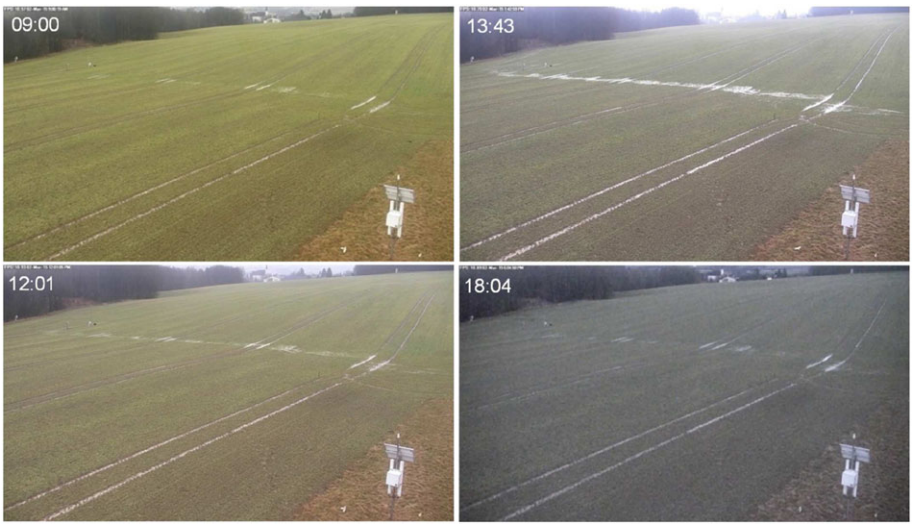

FIGURE 4 Raw time-lapse images during Event 4 (March 2, 2015) with different brightness conditions

was performed by the "Fuzzy Select" tool that selects pixels based on colour similarity. Based on trial and error, we set 30 as the threshold value for "Fuzzy Select," which identified the core regions of wet pixels. The remaining wet pixels that could not be selected by the tool were manually classified using the "Pencil" tool. The time resolution of the sample images used in this manual classification was 15 min.

2.4.3 | Automated classification

The automated classification proposed in this study deals with varying image brightness. Figure 4 shows selected raw pictures of Event 4, illustrating the dramatic change of light conditions during the day, which requires careful attention in the automated classification. We therefore chose to classify the wet pixels using global thresholding

based on grey intensity values as described by Shapiro and Stockman (2001), which only requires a single threshold expressing the brightness conditions instead of three RGB values. Following Shapiro and Stockman (2001), grey intensity was calculated from the pixel RGB components by

$$Y' = 0.2989*R + 0.5870*G + 0.1140*B, \quad (1)$$

where Y′ is grey intensity ranging from 0 (black) to 255 (white) and R, G, and B are the red, green, and blue intensities of a pixel, respectively.

Due to the dynamics of the natural lighting condition, the greyscale images throughout an event have different brightness characteristics, implying the threshold changes between images. A similar problem was encountered by Pawlowicz (2003) in investigating oceanographic phenomena on coastal waters who applied best-fit planes subtraction and a histogram equalization algorithm for large-scale and small-scale brightness variations, respectively, which worked well because the study dealt with a large body of water, thus enough control points could be selected to correct the brightness variations.

In the current study, the wet areas only fill a small part of the image and are especially scarce during the beginning and end of an event, which does not allow a large number of control points for brightness correction. Therefore, instead of correcting the brightness for the entire event, we accommodated brightness variations by introducing a threshold constant C that connects the threshold value T to its greyscale histogram mode M (see Figure 5), an indicator of the image brightness. Provided C is known, the threshold value T of an image can be calculated as

$$T = M + C. \quad (2a)$$

In order to obtain C for the images at 1 min resolution, we calibrated it from a subset of n images (sample images, at 15 min resolution) against the manually classified wet pixels and assumed that C does not vary within an event (but could vary between events). In the calibration procedure, we defined a range of trial constants C_j estimated from trial and error as 0–50. For each sample image, we determined the threshold T based on its mode M and Equation 2a:

$$T_{i,j} = M_i + C_j, \quad (2b)$$

where i = 1…n is the sample image index of n images, j = 1…m is the trial constant index of m trials, $T_{i,j}$ is the threshold value, M_i is the sample image mode, and C_j is the trial threshold constant.

Using this threshold, we classified all n sample images of a given event for different C_j:

$$g_{i,j}(x,y) = \begin{cases} 1, & f(x,y) \geq T_{i,j} \\ 0, & f(x,y) < T_{i,j} \end{cases}, \quad (3)$$

where $g_{i,j}(x,y)$ is a binary value (1 = wet, 0 = dry), $f(x,y)$ is the pixel grey value, $T_{i,j}$ is the trial threshold value (Equation 2b), x and y are the pixel coordinates.

For each C_j, we evaluated the RMSE (m^2) between automated and manual areas of wet pixels as the objective function:

$$\mathrm{RMSE}(C_j) = \sqrt{\frac{\sum_{i=1}^{n} (AA_i(C_j) - AM_i)^2}{n}}, \quad (4)$$

where AA_i (C_j) is the automated area of sample image i classified using C_j, AM_i is the manual area of sample image i, and n is the number of sample images within an event. The areas (AA and AM) are calculated by multiplying the number of wet pixels by the orthoimage pixel area (5 × 5 cm^2).

C_j associated with the minimum RMSE was noted as the optimum threshold parameter C' for each event. C' was then used for

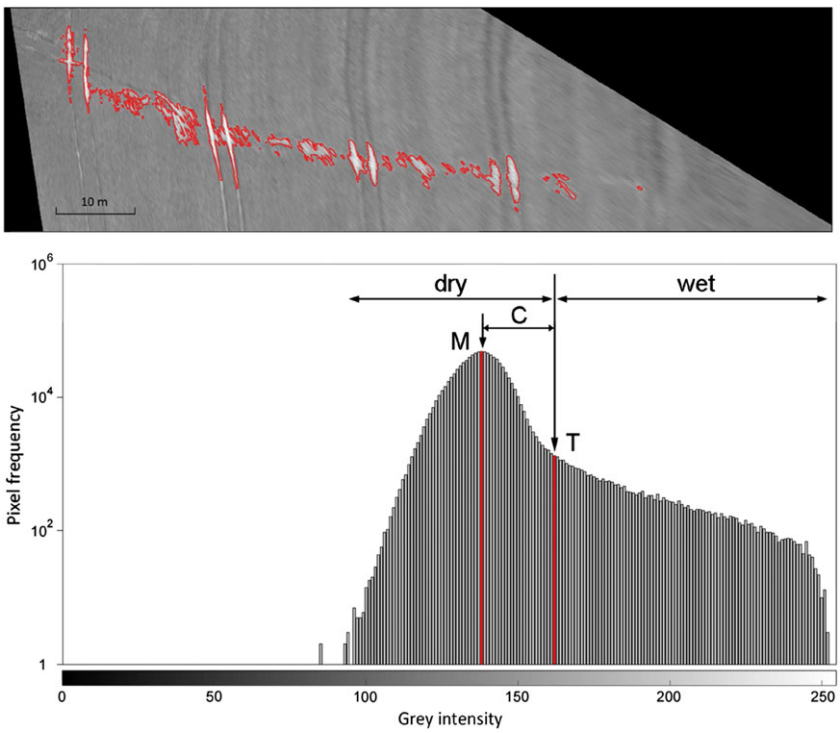

FIGURE 5 Example of greyscale image histogram. Top: Greyscale image sample of event 4 (March 2, 2015, 11:45). The red pixels outline the automatic classified area. Bottom: Greyscale histogram of the top image. Grey intensity ranges from 0 (black) to 255 (white) with blank pixels around the georectified image (black region shown in the top image) excluded. The vertical axis is in log scale. Symbols M, T, and C in the histogram indicate the mode, threshold, and threshold constant of the top image that are 139, 162, and 23, respectively

automatically thresholding all N time-lapse images of the event (at 1 min resolution) in an analogous way as Equation 3 but with T' (calculated using C' and Equation 2a). The area AA (C') was similarly calculated for all N images.

To assess the accuracy of the automated method for different events, we normalized the RMSE by the mean manual area of wet pixels of the event:

$$\mathrm{NRMSE}_k = \frac{\mathrm{RMSE}_k}{\overline{AM_k}}, \tag{5}$$

where k is the event number and RMSE_k and $\overline{AM_k}$ are the root mean squared error and mean manual area of wet pixels of event k, respectively.

2.5 | Accuracy assessment of the spatial patterns

2.5.1 | Comparison with manual classification

The performance of the automated classification was first evaluated by comparing the spatial patterns of wet and dry pixels with those from the manual method following the confusion matrix concept (Provost & Kohavi, 1998; see Table 2). After numerating the correctly classified pixels a (true wet) and d (true dry), and misclassified pixels b (miss, misclassified as dry) and c (false, misclassified as wet), we calculated the accuracy index (K) using the geometric mean, which gives more weight to wet pixels (Kubat, Holte, & Matwin, 1998):

$$K = \sqrt{\left(\frac{a}{a+c}\right)\left(\frac{a}{a+b}\right)}. \tag{6}$$

K = 1 for a perfect match and K = 0 for a complete mismatch.

TABLE 2 Confusion matrix for assessing the accuracy of spatial patterns by comparing the results between automated and manual classifications

		Prediction (automated classification)	
		Wet pixel	Dry pixel
Reference (manual classification)	Wet pixel	a (true wet)	b (miss)
	Dry pixel	c (false)	d (true dry)

2.5.2 | Comparison with field measurements

Additionally, we verified the accuracy of the automated image classification by comparison with field measurements for the event on February 21, 2017. Previous to this event, soil moisture was already high, and rain had started on February 20 in the night. Overland flow was already observed at the start of the camera observations (06:40).

Field measurements were performed by walking around the visible ponds and the overland flow area (see Figure 11 top) and tracing the boundary of the saturation area using the GPS in automatic tracking mode (minute-based point sampling). At around 16:00, the saturation area was well established, which allowed little change during the period of locating the boundary points (about 60 min). We walked around the area six times to ensure the validity of the identified boundary points, during which 61 points were taken.

The error between the automated classification and the field-located boundary points was estimated as the shortest distance between each boundary points of both methods. The error was defined as positive if the field-located boundary point was outside of the automatically classified boundary (implying an underestimation of the automated classification prediction). Negative errors were defined in an analogous way.

3 | RESULTS

3.1 | Observed events

Within the observation period (August 2014–July 2015), there were five overland flow events captured by the camera observations (see Table 3). Because the camera operated in the visible spectrum, night images could not be used. Consequently, Events 1–4 were cut at the end, whereas Event 2 was also slightly cut at the beginning (see the occurrence time in Table 3). The remaining periods did cover the core of the events well.

3.2 | Manual classification

The manual classification of the sample images (sampled every 15 min) was performed for 149 images (Table 4). The lowest saturation area of Event 1 (6.9 m^2, see Table 4) was the area of persistent ponding, which seasonally occurs from autumn to spring where the tractor tracks

TABLE 3 Observed overland flow and rainfall events

		Observable picture time		Overland flow occurrence time			
Event	Date	Start	End	Start	End	Total rainfall (mm/day)	Max. rain intensity (mm/hr)
1	23 Oct 2014	06:49	17:01	12:00	23:59[a]	17.3	2.6
2	11 Jan 2015	07:44	16:53	06:00	23:59[a]	10.2	1.6
3	24 Feb 2015	06:45	17:46	16:40	23:59[a]	6.3	2.4
4	02 Mar 2015	06:31	18:04	09:45	23:59[a]	13.3	4.5
5	11 Mar 2015	06:30	18:18	10:00	12:55[a]	7.0	3.8

[a]Occurrence time continued to the next day.

Note. Time is in UTC + 1. Observable picture time is the period when the camera image is sufficiently bright. Overland flow occurrence time was estimated from the hydrographs at the tile drain outlet (Frau 1).

TABLE 4 Statistics of the saturation area from the manual classification of sample images (sampled every 15 min) for each event

Event	Number of images	Manual saturation Area (m^2)		
		Mean	Min	Max
1	24	62.7	6.9	136.1
2	37	102.4	52.4	153.3
3	21	45.7	24.2	113.2
4	42	74.0	17.1	211.2
5	25	21.9	8.8	34.5
Total	149	66.5	6.9	211.2

intersect with the depression area (see Figure 6). During Events 2–4, rain had occurred since early morning, and the saturation area had already developed for a while before the observable period (in the form of ponding), hence the higher minimum saturation area compared to Events 1 and 5. During late winter, the soil had become wetter, which expanded the persistent ponding spots outside the tractor tracks as illustrated in Figure 6 (taken on February 27, 2015). This condition explains the higher minimum saturation area of Event 5 (March 11) compared to Event 1. The largest saturation area occurred during Event 4, which was 211.2 m^2.

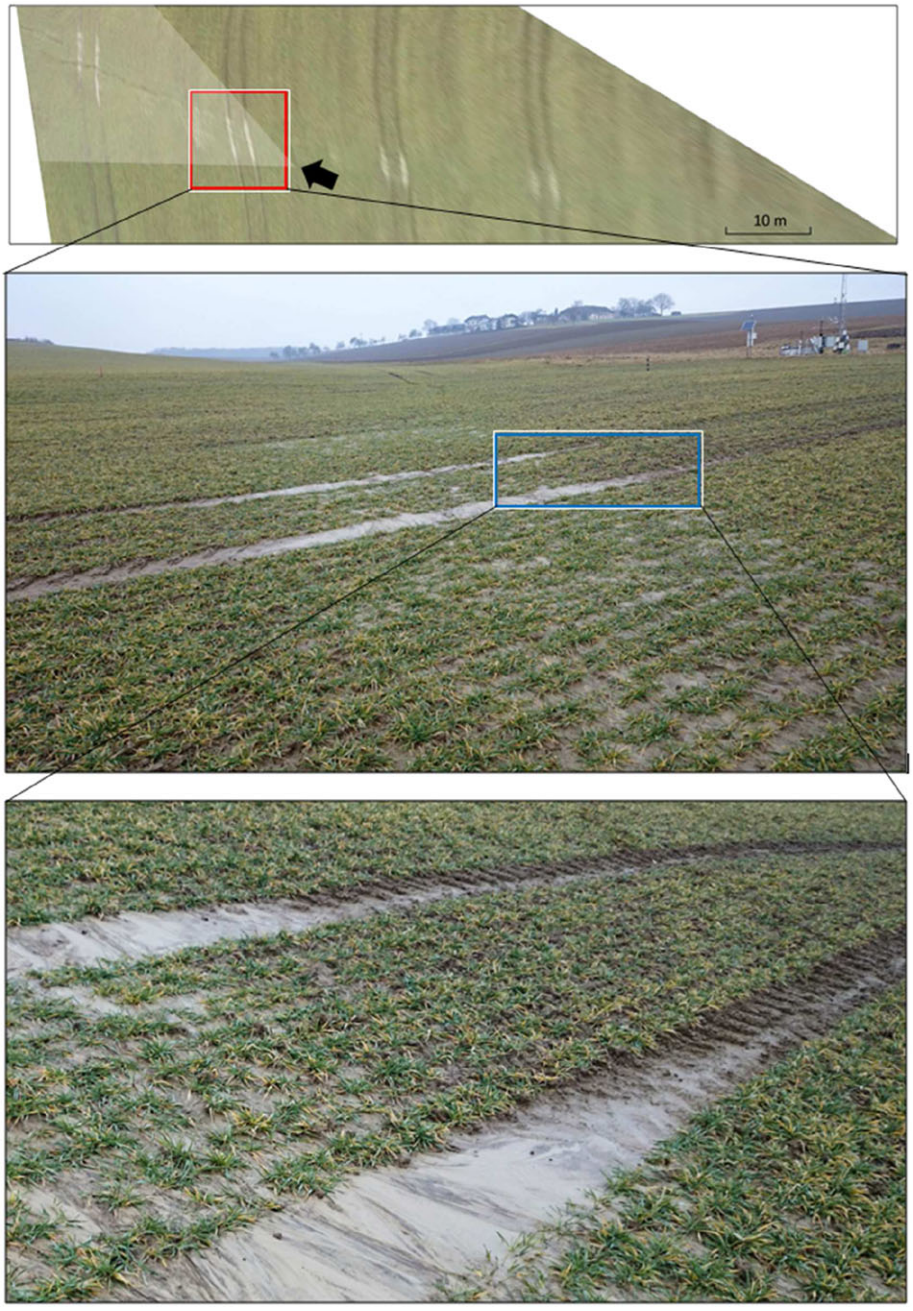

FIGURE 6 Field condition of seasonal ponding on the tractor tracks and the surrounding area. The top panel shows the georectified image from February 27, 2015, 09:40 with the red box highlighting the tractor tracks shown in the middle picture. The black arrow indicates the direction from which the middle picture was taken. The blue box in the middle picture shows the surface conditions along the tractor track highlighted in the bottom picture. Both pictures were taken around the same time as the georectified image

3.3 | Threshold optimization and accuracy of total area

Firstly, threshold T was optimized for each sample image by minimizing the saturation area RMSE between automated and manual results. The optimum T' of the 149 sample images ranged from 148 to 192 and showed a strong correlation with the image mode M (Pearson coefficient 0.91; Figure 7 top left). Linear regression between T' and M (with slope 1) resulted in fitted threshold parameter C = 19, with NRMSE = 21%. Figure 7 (top right) indicates good performance of the automated procedure (NRMSE = 1.5%) regardless of the image brightness defined by M, whereas the individual errors do not correlate with M.

In order to understand the characteristics of the greyscale image histograms, M was correlated with global radiation and local solar time (LST, i.e., the time according to the position of the sun), which influence the lighting condition in terms of sunlight intensity and daylight progression (Figure 7, bottom panels). Global radiation was interpolated from half-hourly measurements at the weather station. LST was calculated based on local time and the site coordinates (Sproul, 2007). The correlations between M and both global radiation and LST are similar (Pearson coefficient −0.21 and −0.20, respectively). Although, during Event 1, global radiation was lower than that of Event 2 (daily averages of 11.5 and 36.6 W/m^2, respectively; Figure 7, bottom left), both events share similar M characteristics (mean and standard deviation 147 and 7.4, and 148 and 8.4 for Events 1 and 2, respectively). This could be due to global radiation being influenced by cloud amount and type (Josefsson & Landelius, 2000); thus, similar sunlight conditions may result in different global radiation values. In the correlation with LST, there is a tendency for higher M (i.e., brighter images) to occur for morning images (Figure 7, bottom right).

When the performance of optimization is evaluated between T- and C-fitting (finding optimum T and C per event, respectively), Figure 8 (top panels) shows that the RMSE of the C-fitting converges faster and in a narrower range, which corresponds to a lower standard deviation of C between events than for T (4.2 and 9.4, respectively), suggesting C as a more robust parameter. For comparison, Table 6 shows the results of applying a single C and T for all events, in which we used C = 19 (linear regression intercept, see Figure 7 top left) and T = 169 (mean T', see Table 5). Overall, C = 19 still gives good results with a lower error than T-fitting (mean NRMSE 21.0% and 37.0%, respectively), although much worse than fitting C individually for each event (mean NRMSE 9.2%). Using single T produced the worst results of all methods (mean NRMSE 44.6%). For the further analysis, event-based C for image classification was used (Table 5, right).

3.4 | Accuracy of the spatial patterns

3.4.1 | Comparison with manual classification

The accuracy K (Equation 6) of automatic classification based on the confusion matrix (Table 2) was first evaluated for each event (Table 7). K varies from 92.6–95.2% with Events 1, 3, and 5 having slightly lower accuracies than the others possibly due to the events' smaller size with less overland flow and lower connectivity between the ponded areas, making the spatial patterns patchier. The patchy patterns tend to produce misclassified pixels around the saturation area boundary (see Figure 9) due to the transition between wet and dry pixels, which creates a mixed signal. Overall, the automated method gives an accuracy of 94.3% as compared to the manual method with 5.4% false and 4.2% miss wet pixels prediction, indicating a very small bias of 1.2% (overestimation of saturation area) relative to the manual classification.

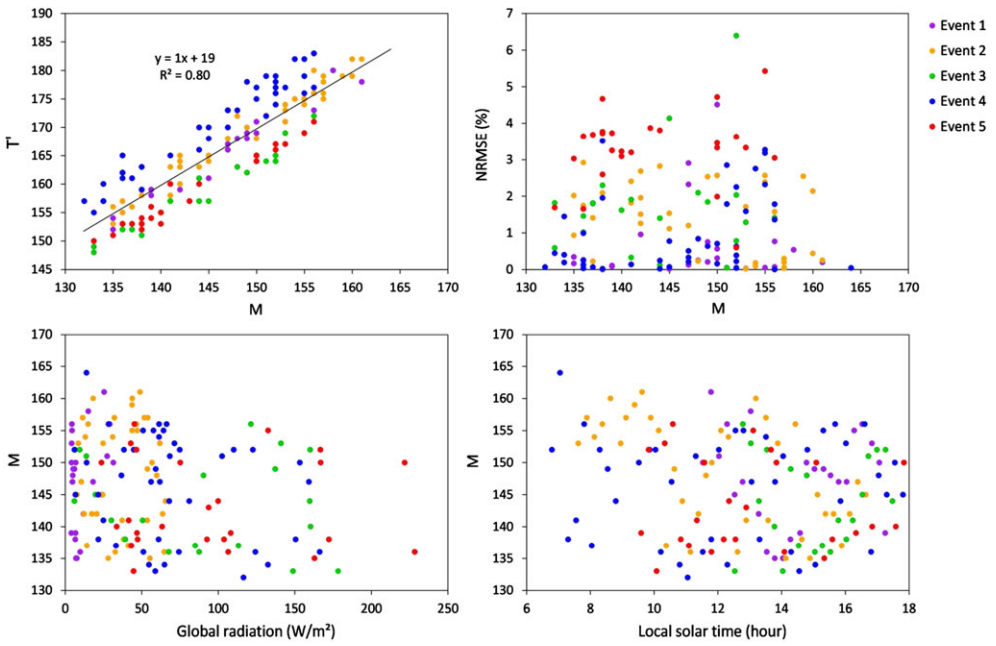

FIGURE 7 Top left: Optimum threshold (T') plotted against image mode (M). Equation represents linear regression. The optimum T' was obtained by minimizing the root mean square error between automated and manual saturation areas for each of the 149 sample images individually. Top right: Normalized root mean square error (NRMSE; Equation 5) plotted against M. Bottom: M plotted against global radiation and local solar time. Each point represents one image

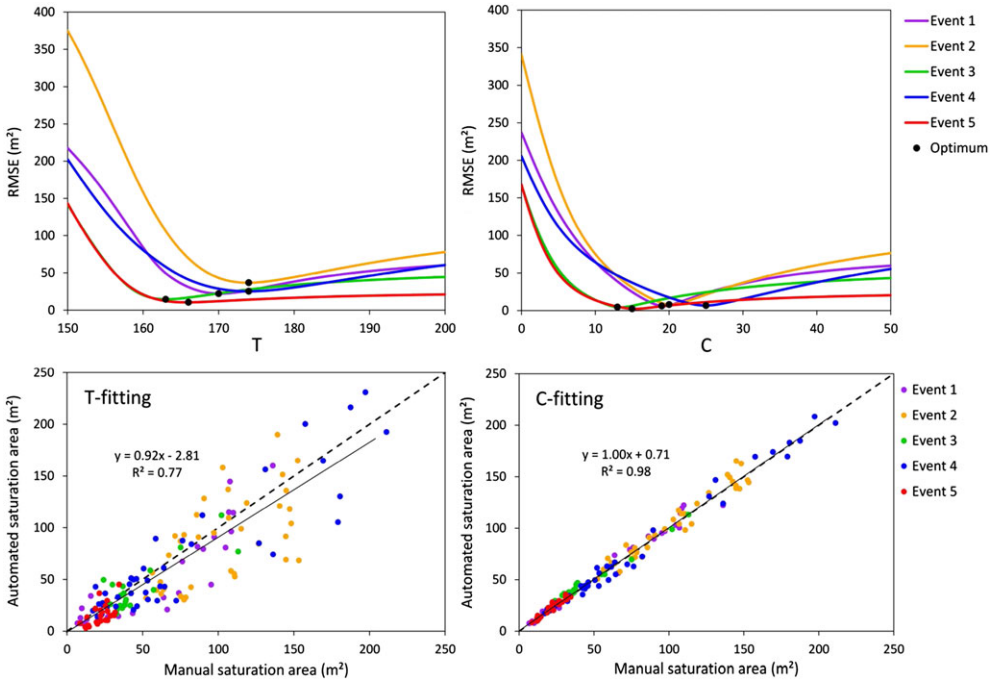

FIGURE 8 Top: Root mean square error (RMSE; Equation 4) of fitting the saturation areas of automated classification against those of the manual classification for a range of thresholds T (150–200) and threshold parameters C (0–50). The black circles indicate the optimum values of T' and C' that yield minimum RMSE for each event (see Table 5). Bottom: Comparison between automated and manual saturation areas of 149 sample images with linear regression (solid line). Dashed line is the 1:1 line. The optimum values of T' and C' obtained from the top panels are used to calculate automated saturation area in the bottom panels

Figure 10 shows the automatic classification accuracy with respect to the distance from the camera. For distances shorter than 90 m, the accuracy is around 95% (Figure 10, bottom left). At longer distances, the accuracy slightly drops as would be expected due to the lower resolution of the georectified images and the flatter camera angle, which both reduce the optical efficiency. For the set-up as used in this study, 100 m distance seems to be the limit at which saturation patterns can be detected with reasonable accuracy. Interestingly, at distance class 55 m, the accuracy is somewhat lower than in the adjacent classes due to the microtopography of the area, which produced patchier patterns (from individual ponds), thus a higher number of misclassified pixels as explained previously.

3.4.2 | Comparison with field measurements

The field observations resulted in 61 points representing the boundary of the saturation area on February 21, 2017, around 16:00 as shown as red points in Figure 11. Although a small section of the boundary could not be accessed as the land had already been cultivated, the points do cover the boundary in a representative way and are thus suitable for validating the automated procedure. The total saturation area identified by the automated classification (blue area in Figure 11) using C = 19 (linear regression intercept, see Figure 7 top left) at this point in time was 269.8 m^2.

Figure 12 shows the errors (shortest distance between measured points and automatically classified boundary) for each distance class from the camera as defined in Figure 10 top. The errors tend to be positive, meaning that there is a small tendency for the method to underestimate the size of the saturation area. The biases are around 9 cm (Table 8). This bias is related to the inaccuracy in estimating the saturation area at the boundary where a thin water film was visible, which the photograph failed to capture. At the shortest distance class of 45 m, the median error is negative (Figure 12). but the bias is still

TABLE 5 Saturation areas from the automated classification using calibration of the optimum threshold (T') and the threshold parameter (C') individually for each event, along with associated root mean square errors as compared to the manual saturation areas

	T-fitting				*C*-fitting			
Event	T'	Mean Area (m^2)	RMSE (m^2)	NRMSE (%)	C'	Mean Area (m^2)	RMSE (m^2)	NRMSE (%)
1	170	54.6	22.1	35.3	19	63.4	6.0	9.5
2	174	87.9	36.7	35.8	20	103.6	7.9	7.7
3	163	39.8	14.6	31.8	13	48.4	4.5	9.8
4	174	68.7	25.2	34.0	25	73.5	6.6	8.9
5	166	15.2	10.5	48.0	15	22.4	2.2	10.2
Average	169	53.3	21.8	37.0	18	62.3	5.4	9.2

Note. Root mean square error (RMSE) is calculated between automated and manual saturation areas for each event. NRMSE is RMSE normalized by mean manual saturation area per event (see Table 4).

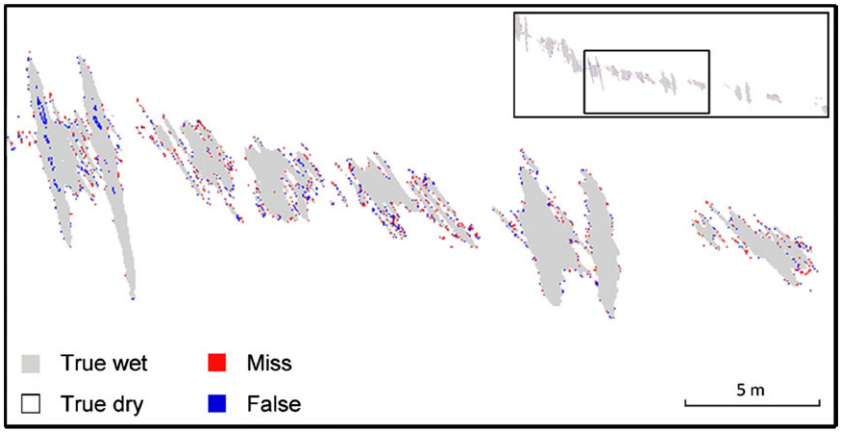

FIGURE 9 Example of accuracy pattern between automated and manual saturation patterns. The main image represents the boxed section in the inset. The grey, white, red, and blue pixels indicate true wet, true dry, miss (misclassified as dry), and false (misclassified as wet), see Table 2. The full image in the inset is part of Event 1 (October 23, 2014, 15:45) with accuracy index of K = 95.9%

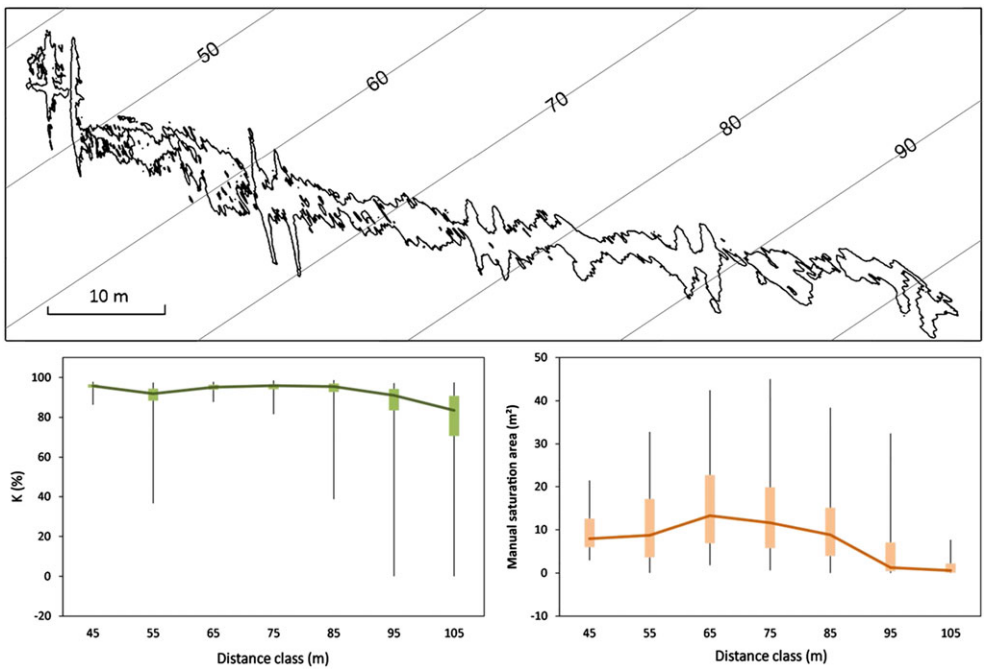

FIGURE 10 Top: Thin lines show pixel distance from the camera (in metre). Thick lines show the maximum expansion of observed saturation area. Area of distance class 45 m, 55 m, and so on used in the bottom panels is defined between distance lines 40–50 m, 50–60 m, and so on. Bottom: Accuracy K (Equation 6) of the spatial patterns of automated classification (left) and saturation areas from manual classification (right) for the 149 sample images (i.e., all events) at each distance class. Box whisker plots and bold lines show the quantiles (minimum, 25th, 75th, and maximum) and median of variability between sample images, respectively

slightly positive (+4 cm, Table 8), and the errors do not decrease as they did in the comparison between automated and manual classification (Figure 10). This is related to the patchier saturation patterns in this part of the domain, which is difficult to identify in detail in the field. The automatic method can more easily trace the patchy patterns. The RMSE over all distance classes is 23 cm (around 5 pixels) with a mean absolute error of 18 cm. These results suggest an excellent performance of the method given the accuracies of the georectification and the GPS.

3.5 | Example of saturation area dynamics during events

Figure 13 shows the saturation area dynamics of Event 4 (March 2, 2015) from automatic classification of 593 images within the observable time (06:31 to 18:04) plotted with tile drain outflow (Frau1) and cumulative rainfall. Spatial patterns of the saturation area at different stages during the event are shown in Figure 14. From this information, we were able to identify three phases of saturation area development and the underlying processes: ponding, run on, and rainfall on run on (shaded areas in Figure 13).

The saturation area graph of Event 4 starts at 6:31 at 20 m^2, which is much larger than the initial seasonal saturation area (6.9 m^2, see Table 4), due to early morning rainfall forming small ponds in the upstream part of the thalweg (Figure 14a). Formation of small ponds continued until 12:00, a period that we termed "ponding" phase. During the ponding phase, individual disconnected small ponds formed due to rainwater that could not infiltrate into the saturated soil as can be seen in Figure 14b. The ponding mechanism is caused by saturation excess, which can be inferred from the soil moisture measurements that showed saturated values during the time of Event 4.

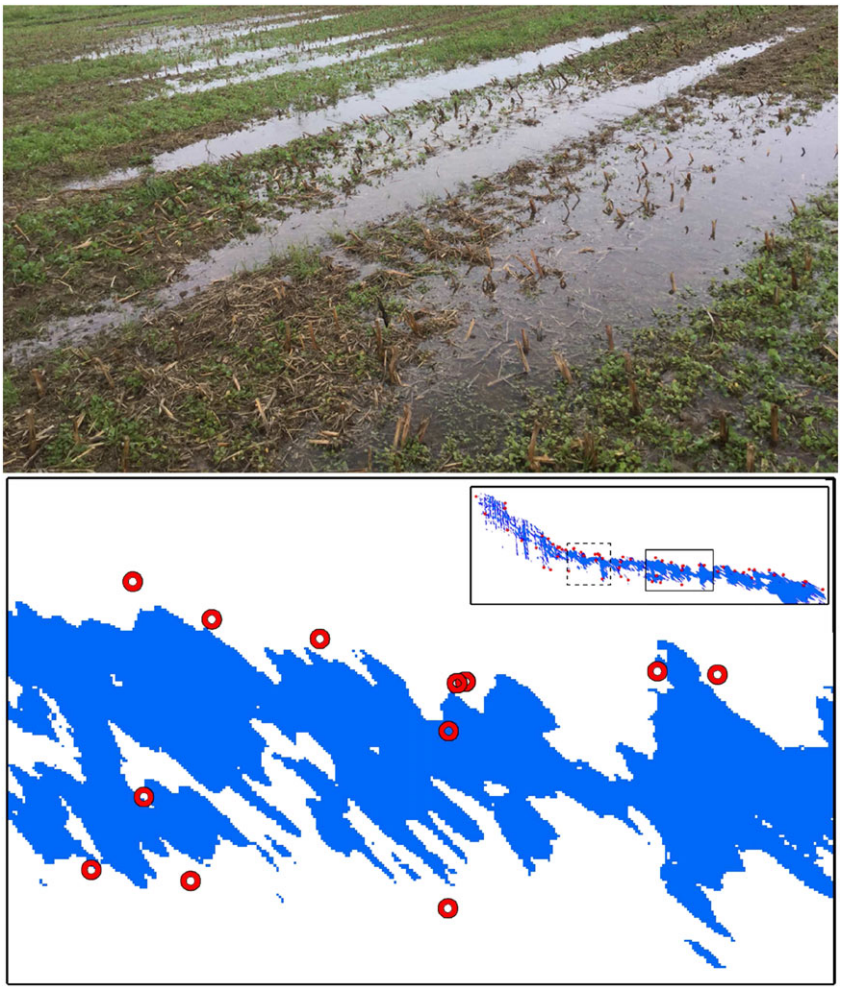

FIGURE 11 Top: Visible saturation area (ponding and overland flow) during the field survey. Bottom: Comparison of field measurements of the boundary of the saturation area (red points) and saturation area estimated by the automatic classification (blue area) for event of February 21, 2017, at 16:00. The pixel size is 5 cm. The inner circle diameter of the red points corresponds to the GPS positioning root mean square error of 10 cm. The dashed and solid boxes in the inset mark the area coverage of top and bottom images, respectively

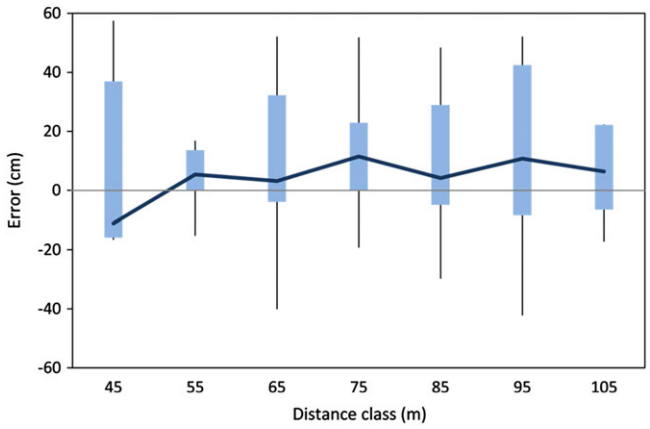

FIGURE 12 Errors (shortest distance) between the boundary of the saturation area estimated by the automated classification and 61 points located during the field survey representing that boundary on February 21, 2017, around 16:00. Positive errors correspond to field-located points outside of the classified boundary. Box whisker plots and bold lines show the quantiles (minimum, 25th, 75th, and maximum) and median of variability between measured points at each distance class, respectively. The thin grey line represents y = 0

The second phase is "run on," which occurred from 12:01 to 12:31. This phase was dominated by overland flow from the upstream area of the catchment. During field visits, we observed that surface run-off formed a 30-cm-wide rill channel along the thalweg and connected the individual ponds it passed. As the run on flowed through the ponds, it expanded the saturation area. Because the run on water also spilled to the adjacent area, new saturation area extensively developed around the rill channel and reached the farther ponds, thus increasing the saturation area expansion rate rapidly. The saturation area increased by 110 m^2 in half an hour (from 12:01 to 12:31) compared to only 40 m^2 in five and a half hours in the previous phase (from 06:31 to 12:00). The fast saturation area development during this phase can be seen in Figure 14c–d. There is a time lag between the run-off formation on the saturated upstream area and run-off arrival at the foot of the hillslope as indicated by the less steep increase of the tile drain outflow as compared to the expansion of the saturation areas.

The last phase of saturation area development is "rainfall on run on," which occurred from 12:32 until 13:43 when the saturation area started to shrink. During this phase, most of the individual ponds at the centre of the thalweg area from the ponding phase were already connected with each other. Figure 13 shows that the rainfall intensity was high during this period, which substantially increased tile drain outflow. Because most of the ponds were connected, much of the overland flow occurred as sheet flow, although with different flow depths due to the microtopography. The rate of saturation area expansion is lower than in the previous phase as most of the expansion occurred along the boundary of the saturation area (Figure 14e–f) from ponds in the thalweg core area. Although the expansion rate declined,

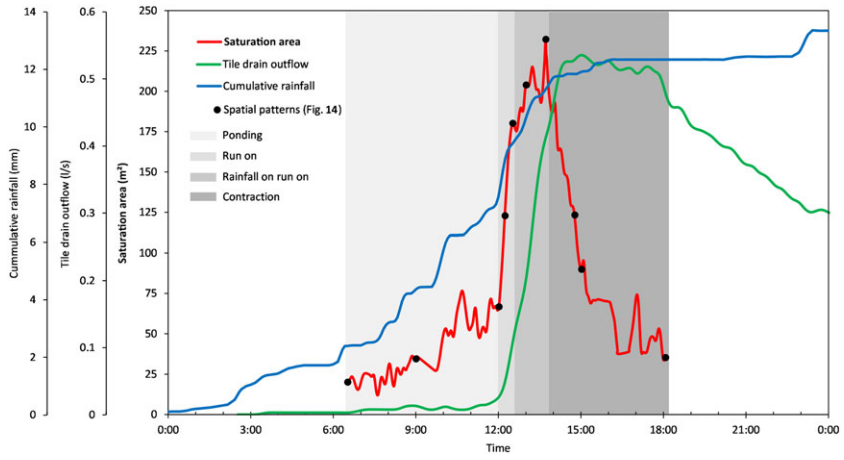

FIGURE 13 Example of saturation area dynamics of Event 4 (March 2, 2015). Red, green, and blue lines represent saturation area (from automated classification), tile drain outflow, and cumulative rainfall, respectively. The black circles relate to the spatial patterns in Figure 14. The grey shadings indicate three phases of the expansion of saturation area and the contraction phase. Saturation area dynamics are plotted within the observable time (06:31–18:04). Tile drain outflow and cumulative rainfall are aggregated from 1 min data to 5 min

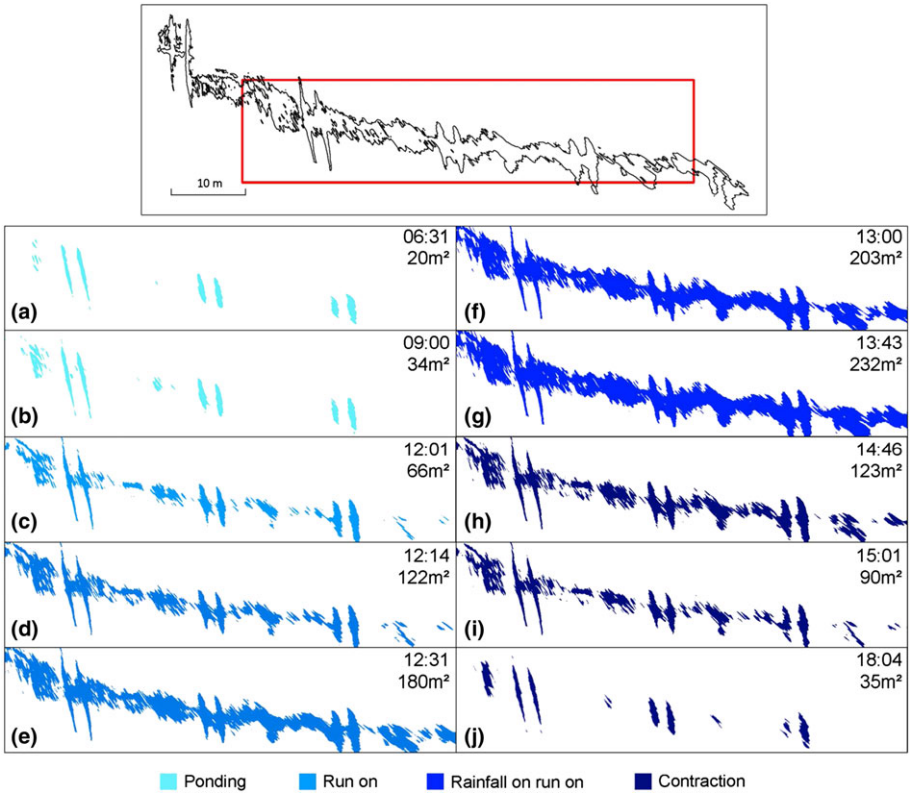

FIGURE 14 Spatial patterns of saturation dynamics during Event 4 (March 2, 2015) based on automated classification. Black outline in the top panel indicates the pattern of maximum saturation area at 13:43. The blue patterns represent the saturation area inside the red box shown in the top panel with different shades based on the respective phase. Numbers are image time and automated total saturation area (*AA*). Image times have been chosen irregularly to illustrate the key phases of the saturation area development (grey shadings in Figure 13). Figure 4 shows the raw images of Figure 14b,c,g,j

it did not affect the increasing rate of tile drain flow. This is possibly due to the insignificance of saturation area at the observed thalweg (0.02 ha, maximum observed saturation area) compared to the upstream area (5 ha, estimated from the DEM). The contracting saturation area at the thalweg did not affect the increase of run on flow, which mostly came from the upstream area.

As the rainfall intensity decreased and rainfall eventually stopped, the saturation area contracted quickly and disconnected within an hour (Figure 14h–i). Three hours later, the patterns (Figure 14j) and the extent were very similar to those at the beginning of the observation period (Figure 14a).

4 | DISCUSSION

4.1 | Automated classification method for mapping saturation area

We propose an automatic classification method with thresholding based on grey intensities instead of the RGB values of previous studies (e.g., Keys et al., 2016). This choice has been motivated by (a) simplifying the threshold assignment for a large number of images by reducing three thresholds (RGB) to a single threshold (i.e., grey intensity) and (b) directly adapting to the highly varying image

brightnesses found in this study. Because high-frequency time-lapse photograph observations produce a large number of images, automated procedures are preferable to manual analyses that commonly use threshold values fitted by visual inspection (Lang, Shao, Pijanowski, & Farnsworth, 2008). We calibrated the automated results against manual analyses (149 sample images) and determined the optimum threshold for each sample image. Although there are concerns regarding the subjectivity of manual classification (Smits, Dellepiane, & Schowengerdt, 1999; Zha, Gao, & Ni, 2003), it has also been recommended for complementing the automated procedures to increase the accuracy when misclassification is likely due to mixed signatures of different classes (Härer, Bernhardt, Corripio, & Schulz, 2013). These may occur if cloud shadows and sun reflections affect the contrast between saturation area and its surrounding.

The automatic classification thresholds optimized against manual results were found to be linearly correlated with the image greyscale mode *M* (Figure 7 top left), suggesting that thresholding based on each picture's *M* may be viable. Therefore, we introduced a threshold parameter *C* as the difference between threshold *T* and *M*, which was calibrated to the manual results (15 min sampling). A slope 1 regression between *T* and *M* for all images (Figure 7 top left) gives an intercept of *C* = 19, which could be used to estimate *T* for all images. However, *C* can also be fitted per event to increase the accuracy. Comparisons between constant *C* = 19 and fitted-*C* per event gave NRMSE of 21% and 9.2%, respectively, (Tables 6 and 5) with $R = 0.98$ (Figure 8 bottom right). This error is comparable with the results of Keys et al. (2016) based on three (RGB) thresholds, with NRMSE = 7.96% and $R^2 = 0.94$.

The advantage of this automated method is its ability to handle varying image brightnesses within a large number of images produced from high-frequency observations (Figure 4). Keys et al. (2016) dealt with varying image brightnesses by choosing different threshold values for each image, but they analysed hourly observations. Alternatively, Pawlowicz (2003) applied image corrections using best-fit planes subtraction and histogram equalization requiring extensive image histogram analysis for each type of image brightness, a nonviable option in this study. Another possibility would be to complement the observations with a thermal (IR) camera of compatible pixel resolutions, as IR images are not affected by the lighting conditions (although varying sunlight intensity on the surface needs to be considered in the pixel temperature calibration, which is important in observing agricultural land).

TABLE 6 Saturation areas from the automated classification using calibration of the optimum threshold (*T′*) and the threshold parameter (*C′*) jointly for all events, along with associated root mean square errors as compared to the manual saturation areas

	T′ = 169			*C′* = 19		
Event	Mean Area (m^2)	RMSE (m^2)	NRMSE (%)	Mean Area (m^2)	RMSE (m^2)	NRMSE (%)
1	58.2	22.7	36.2	63.4	6.0	9.6
2	113.1	47.7	46.6	108.4	10.0	9.8
3	28.7	21.0	45.9	34.9	14.7	32.1
4	83.4	30.9	41.7	91.6	20.8	28.1
5	12.1	11.5	52.6	16.6	5.6	25.8
Average	59.1	26.8	44.6	63.0	11.4	21.0

Note. Root mean square error (RMSE) is calculated between automated and manual saturation areas for each event. NRMSE is RMSE normalized by mean manual saturation area per event (see Table 4).

The automated method estimates the spatial saturation patterns with a mean accuracy of 94.3% (Table 7). relative to the manual patterns, which is higher than the 82% accuracy of the object-based classification of satellite images by Dribault et al. (2012). Although the cases are not directly comparable, they do point towards a good performance of the simpler method proposed here. Misclassified pixels in this study mostly occur around the saturation area boundary (Figure 9). which is likely related to mixed signatures. The accuracy of the patterns prediction decreases with distance from the camera (Figure 10 bottom), which is related to the lower resolution of the rectified image and the flatter angle of the camera with respect to the terrain, a particular problem in highly oblique photographs (Bourgault, 2008). The accuracy could be improved by increasing the observation angle and/or using a higher resolution camera. An example of the latter is Härer et al. (2013) who classified snow cover at 900 m distance using a 17.9 megapixels SLR camera (nine times the present network camera resolution) with 0.9% misclassification.

To assess the validity of the automatic method, comparisons to ground truth from field surveys are advised (Tiner, Mulamoottil, Warner, & McBean, 1996). We collected field measurements for an event independent of the remaining analysis (February 21, 2017) by walking along the saturation area boundary and taking GPS positions at 61 points. The error was defined as the shortest distance between the measured points and the automatically classified boundary (using *C* = 19), which resulted in an RMSE = 23 cm (Table 8). This error consists of three components: (a) classification error incurred by potential misclassification at the right location, (b) image georectification error from an inaccurate DEM, and (c) GPS positioning error of the GCPs. The georectification RMSE was estimated as 17 cm (see Section 2.4.1). The GPS positioning RMSE was given by the manufacturer of the device as 10 cm. Assuming that the saturation areas from the field survey represent the ground truth and the errors are independent, this would suggest a classification error of 12 cm ($RMSE_{class}^2 = RMSE_{total}^2 - RMSE_{georect}^2 - RMSE_{gps}^2$). This error is considered suitable for meaningfully delineating saturation areas at the hillslope scale. The bias of the automated method against the field measurements was around +9 cm, implying an underestimation of the saturation area. The length of the saturation area boundary was 240 m, which translates into an areal bias of 240 * 0.09 = 21.6 m^2 or 8% of the total area (269.8 m^2).

Field mapping is time consuming and thus not suitable for fine temporal resolutions (e.g. under 30 min). Previous field studies therefore produced rather low resolutions, both in space and time (Dunne & Black, 1970; Blazkova et al., 2002; Meerkerk et al., 2009). The proposed time-lapse photography, paired with automatic image classification, overcomes this limitation and provides detailed insight into run-off generation mechanisms.

Although the proposed method has commendable strengths, it does have difficulties during two conditions: dark conditions, and when the crop obstructs the ground. The first issue could be addressed by complementary IR imagery observations (Deitchman & Loheide, 2009; Pfister et al., 2010). Although the spatial resolution of previous

TABLE 7 Accuracy of the automated spatial pattern predictions as compared to manual patterns for the 149 sample images

Event	Average number of predicted pixels				Prediction (%)			
	a (true wet)	b (miss)	c (false)	$p = a + b + c$ (total prediction)	True wet	Miss	False	*K*
1	23,536.8	1,195.8	1,486.0	26,218.6	89.8	4.6	5.7	93.8
2	38,674.2	1,661.6	2,156.5	42,492.4	91.0	3.9	5.1	95.2
3	17,494.2	542.8	1,614.6	19,651.6	89.0	2.8	8.2	93.7
4	27,753.1	1,440.7	1,269.1	30,462.9	91.1	4.7	4.2	95.0
5	8,059.7	516.4	741.4	9,317.5	86.5	5.5	8.0	92.6
Average	25,035.8	1,174.5	1,484.6	27,694.8	90.4	4.2	5.4	94.3

Note. Average number of predicted pixels (automated versus manual method) and average prediction percentage for true wet, miss (misclassified as dry), and false (misclassified as wet) pixels for each event and average over all events.

K is geometric mean accuracy (Equation 6) in percent. Predictions of true wet, miss, and false pixels are calculated from *a*, *b*, and *c* divided by *p* (total prediction) in percent, respectively.

TABLE 8 Statistics of errors (shortest distance) between the boundary of the saturation area estimated by the automated classification and 61 points located during the field survey representing that boundary on February 21, 2017, around 16:00

Distance class	Number of measured points	Error (cm)			
		Bias	Std. dev.	RMSE	Mean absolute error
45	5	4	23	24	21
55	6	5	11	12	10
65	17	9	22	24	18
75	10	16	21	26	20
85	11	8	22	23	19
95	8	9	28	29	25
105	4	3	14	14	13
All	61	9	22	23	18

Note. Positive errors correspond to field-located points outside of the classified boundary. Bias is the mean of the errors, Std. dev. is the standard deviation, RMSE is the root mean squared error, and the mean absolute error is the mean of the absolute values of the errors.

studies was lower (< 1 Megapixels, observed area < 10 m^2), technology improvement may make thermal cameras an attractive option in the future. The second issue of crop obstruction could be addressed by an additional camera with a steeper viewing direction. Alternatively, soil wetness can be observed using a network of soil moisture sensors (Blöschl et al., 2016), whereas overland flow occurrence can be detected by overland flow detectors (e.g. Vertessy & Elsenbeer, 1999; Moody & Martin, 2015).

4.2 | Potential of time-lapse photography in understanding hydrological processes

The example of observed spatial dynamics of saturation areas presented in Figure 14 illustrates the capability of the proposed method to provide two main pieces of information at the hillslope scale: (a) surface run-off generation phases during natural rainfall events (i.e. from ponding, to run on in the form of rill flow, to developed overland flow in the form of sheet flow, to contraction), and (b) the dynamics of spatial patterns and connectivity of saturation areas (i.e. from patchy ponds, to connected ponds by run on, to fully connected overland flow, to disconnection during contraction). Analysing these pieces of information together with other hydrological data, such as rainfall and discharge as we did in Figure 13, can provide deeper insights into surface run-off processes in the context of a number of hydrological science questions.

First, the main benefit of visual observations is providing information on the spatial patterns and connectivity which cannot be obtained from hydrographs or soil moisture measurements at individual locations. Although variable source areas (VSA) have been found to include both near-stream and upstream areas (McDonnell, 2003; Jencso & McGlynn, 2011), the way the source areas spatially connect and form preferential surface flow paths is still in discussion and mostly studied through hydrograph and soil moisture data, and through simulations (Western et al., 2001; Jencso et al., 2009; Bachmair & Weiler, 2014). A stronger focus on visual observations will clearly shed light on how source areas develop from patchy small areas to a connected large area, which acts as both a significant producer and a preferential flow path of surface run-off.

Second, the observed saturation areas inferred from the photographs can assist in distinguishing the main mechanisms of surface run-off generation, in particular, infiltration and saturation excess. These mechanisms are usually studied by numerical and laboratory models (Gevaert et al., 2014; Peñuela et al., 2015). The availability of dynamic data at a meaningful spatial scale such as hillslopes promises significant progress. Pairing time-lapse photography with subsurface hydrological data has the potential of correlating surface and subsurface flow to gain insight into the mechanisms of saturation area development.

Third, information about spatial patterns can improve spatially distributed model testing in the spirit of Grayson and Blöschl (2001). Although a lot of interest has been placed in how saturation areas develop and contribute to surface run-off, most studies applied numerical models and tested them only against stream hydrograph data at the outlet (Mirus, Ebel, Heppner, & Loague, 2011; Weill et al., 2013). As more sophisticated numerical models are being developed, especially to capture the spatial connectivity of saturation areas (Yang & Chu, 2013; Appels, Bogaart, & van der Zee, 2016), their testing against observed spatial patterns of the hydrological dynamics becomes even more important. The observed patterns will assist in ensuring that not only correct hydrographs at the catchment outlet

are simulated but also realistic internal processes within the catchment (Tetzlaff, Carey, Laudon, & McGuire, 2010).

Finally, the method could be extended to the catchment scale by adding cameras to include other hillslopes, and the patterns could be stitched together to cover the entire catchment area. For such an exercise, higher camera resolutions would be advisable to increase the accuracy and the spatial resolution. When the entire source area is observed, the information can be used to estimate catchment scale surface run-off contributing area in a dynamic way. Ultimately, the focus on patterns will assist in making progress in the upscaling problem that has haunted hydrologists since the science began (Blöschl, 2006).

5 | CONCLUSIONS

We propose a new method for observing spatial saturation patterns with high spatial and temporal resolution (5 cm pixels at 1 min intervals). The results demonstrate the potential of time-lapse photography to provide detailed insight into the run-off generation processes and the fine scale connectivity behaviour of surface run-off during natural event. Overall, we conclude as follows:

- Automatic classification of saturation areas using grey intensity thresholds is able to deal with highly variable image brightnesses of time-lapse photography at 1 min resolution by introducing a threshold parameter *C*, which adapts the optimum threshold *T* to the brightnesses of individual images.
- A comparison of the automatic classification with a manual classification gives good performance in terms of both saturation area estimation (mean NRMSE = 9.2%) and pattern estimation (mean accuracy K = 94%) when *C* is calibrated per event. Verification against field measurements for one date as a preliminary demonstration shows very good agreement with an RMSE = 23 cm and a bias = 9 cm.
- The dynamics of the saturation patterns captured by the high temporal resolution provide insights into surface run-off processes that cannot be obtained by hydrograph measurements. The method therefore holds substantial promise for furthering the understanding of hydrological connectivity, assisting in upscaling methods, distinguishing run-off generation mechanisms, estimating run-off contributing area, and testing spatially distributed models.

Future work will explore saturation patterns for a broader spectrum of catchment processes to provide a more in-depth analysis of the phases of saturation area development and to quantify the spatial connectivity of saturation patterns and their effects on run-off response.

ACKNOWLEDGMENTS

The authors would like to thank the Austrian Science Fund (FWF) for the funding as part of the Vienna Doctoral Programme on Water Resource Systems (DK W1219-N22). We would also like to thank Markus Oismüller and Matthias Oismüller for assistance during field work; Thomas Bauer, Stefan Härer, and Prof. Josef Jansa for discussions on image processing; and the staff of the Institute for Land and Water Management Research (BAW-IKT) at Petzenkirchen for their constant assistance.

REFERENCES

Allard, D., & Group, H. (1993). On the connectivity of two random set models: The truncated Gaussian and the Boolean. In A. Soares (Ed.), *Geostatistics Tróia '92* (pp. 467–478). Dordrecht, Netherlands: Springer. https://doi.org/10.1007/978-94-011-1739-5_37

de Alwis, D. A., Easton, Z. M., Dahlke, H. E., Philpot, W. D., & Steenhuis, T. S. (2007). Unsupervised classification of saturated areas using a time series of remotely sensed images. *Hydrology and Earth System Sciences*, *11*(5), 1609–1620. https://doi.org/10.5194/hess-11-1609-2007

Antoine, M., Javaux, M., & Bielders, C. (2009). What indicators can capture runoff-relevant connectivity properties of the micro-topography at the plot scale? *Advances in Water Resources*, *32*(8), 1297–1310. https://doi.org/10.1016/j.advwatres.2009.05.006

Appels, W. M., Bogaart, P. W., & van der Zee, S. E. A. T. M. (2016). Surface runoff in flat terrain: How field topography and runoff generating processes control hydrological connectivity. *Journal of Hydrology*, *534*, 493–504. https://doi.org/10.1016/j.jhydrol.2016.01.021

Aschenwald, J., Leichter, K., Tasser, E., & Tappeiner, U. (2001). Spatio-temporal landscape analysis in mountainous terrain by means of small format photography: A methodological approach. *IEEE Transactions on Geoscience and Remote Sensing*, *39*(4), 885–893. https://doi.org/10.1109/36.917917

Bachmair, S., & Weiler, M. (2014). Interactions and connectivity between runoff generation processes of different spatial scales. *Hydrological Processes*, *28*(4), 1916–1930. https://doi.org/10.1002/hyp.9705

Barrette, J., August, P., & Golet, F. (2000). Accuracy assessment of wetland boundary delineation using aerial photography and digital orthophotography. *Photogrammetric Engineering & Remote Sensing*, *66*(4), 409–416.

Bauer, T., & Strauss, P. (2014). A rule-based image analysis approach for calculating residues and vegetation cover under field conditions. *Catena*, *113*, 363–369. https://doi.org/10.1016/j.catena.2013.08.022

Blazkova, S., Beven, K. J., & Kulasova, A. (2002). On constraining TOPMODEL hydrograph simulations using partial saturated area information. *Hydrological Processes*, *16*(2), 441–458. https://doi.org/10.1002/hyp.331

Blöschl, G. (2006). Hydrologic synthesis: Across processes, places, and scales. *Water Resources Research*, *42*(3), n/a–n/a). https://doi.org/10.1029/2005WR004319

Blöschl, G., Blaschke, A. P., Broer, M., Bucher, C., Carr, G., Chen, X., … Zessner, M. (2016). The Hydrological Open Air Laboratory (HOAL) in Petzenkirchen: A hypothesis-driven observatory. *Hydrology and Earth System Sciences*, *20*(1), 227–255. https://doi.org/10.5194/hess-20-227-2016

Blöschl, G., & Sivapalan, M. (1995). Scale issues in hydrological modelling: A review. *Hydrological Processes*, *9*(3–4), 251–290. https://doi.org/10.1002/hyp.3360090305

Blume, T., & van Meerveld, H. J. (Ilja)(2015). From hillslope to stream: Methods to investigate subsurface connectivity. *Wiley Interdisciplinary Reviews: Water*, *2*(3), 177–198. https://doi.org/10.1002/wat2.1071

Bourgault, D. (2008). Shore-based photogrammetry of river ice. *Canadian Journal of Civil Engineering*, *35*(1), 80–86. https://doi.org/10.1139/L07-087

Bracken, L. J., & Croke, J. (2007). The concept of hydrological connectivity and its contribution to understanding runoff-dominated geomorphic systems. *Hydrological Processes*, *21*(13), 1749–1763. https://doi.org/10.1002/hyp.6313

Bracken, L. J., Wainwright, J., Ali, G. A., Tetzlaff, D., Smith, M. W., Reaney, S. M., & Roy, A. G. (2013). Concepts of hydrological connectivity: Research approaches, pathways and future agendas. *Earth-Science Reviews*, *119*, 17–34. https://doi.org/10.1016/j.earscirev.2013.02.001

Brun, C., Bernard, R., Vidal-Madjar, D., Gascuel-Odoux, C., Merot, P., Duchesne, J., & Nicolas, H. (1990). Mapping saturated areas with a helicopter-borne C band scatterometer. *Water Resources Research*, *26*(5), 945–955. https://doi.org/10.1029/WR026i005p00945

Deitchman, R. S., & Loheide, S. P. (2009). Ground-based thermal imaging of groundwater flow processes at the seepage face. *Geophysical Research Letters*, *36*(14), n/a–n/a). https://doi.org/10.1029/2009GL038103

Dobos, E. (2006). Albedo. In R. Lal (Ed.), *Encyclopedia of soil science*. London, England: Taylor & Francis. Retrieved from http://www.uni-miskolc.hu/~ecodobos/14334.pdf

Dribault, Y., Chokmani, K., & Bernier, M. (2012). Monitoring seasonal hydrological dynamics of minerotrophic peatlands using multi-date GeoEye-1 very high resolution imagery and object-based classification. *Remote Sensing*, *4*(7), 1887–1912. https://doi.org/10.3390/rs4071887

Dunne, T., & Black, R. D. (1970). An experimental investigation of runoff production in permeable soils. *Water Resources Research*, *6*(2), 478–490. https://doi.org/10.1029/WR006i002p00478

Eder, A., Strauss, P., Krueger, T., & Quinton, J. N. (2010). Comparative calculation of suspended sediment loads with respect to hysteresis effects (in the Petzenkirchen catchment, Austria). *Journal of Hydrology*, *389*(1–2), 168–176. https://doi.org/10.1016/j.jhydrol.2010.05.043

Exner-Kittridge, M., Strauss, P., Blöschl, G., Eder, A., Saracevic, E., & Zessner, M. (2016). The seasonal dynamics of the stream sources and input flow paths of water and nitrogen of an Austrian headwater agricultural catchment. *Science of the Total Environment*, *542*, *Part A*, 935–945. https://doi.org/10.1016/j.scitotenv.2015.10.151

Gevaert, A. I., Teuling, A. J., Uijlenhoet, R., DeLong, S. B., Huxman, T. E., Pangle, L. A., … Troch, P. A. (2014). Hillslope-scale experiment demonstrates the role of convergence during two-step saturation. *Hydrology and Earth System Sciences*, *18*(9), 3681–3692. https://doi.org/10.5194/hess-18-3681-2014

Glaser, B., Klaus, J., Frei, S., Frentress, J., Pfister, L., & Hopp, L. (2016). On the value of surface saturated area dynamics mapped with thermal infrared imagery for modeling the hillslope-riparian-stream continuum. *Water Resources Research*, *52*(10), 8317–8342. https://doi.org/10.1002/2015WR018414

Grayson, R., & Blöschl, G. (2001). *Spatial patterns in catchment hydrology: Observations and modelling*. Cambridge, England: Cambridge University Press.

Grayson, R. B., Blöschl, G., Western, A. W., & McMahon, T. A. (2002). Advances in the use of observed spatial patterns of catchment hydrological response. *Advances in Water Resources*, *25*(8–12), 1313–1334. https://doi.org/10.1016/S0309-1708(02)00060-X

Güntner, A., Seibert, J., & Uhlenbrook, S. (2004). Modeling spatial patterns of saturated areas: An evaluation of different terrain indices. *Water Resources Research*, *40*(5), W05114. https://doi.org/10.1029/2003WR002864

Härer, S., Bernhardt, M., Corripio, J. G., & Schulz, K. (2013). PRACTISE—Photo Rectification And ClassificaTIon SoftwarE (V.1.0). *Geoscientific Model Development*, *6*(3), 837–848. https://doi.org/10.5194/gmd-6-837-2013

Jencso, K. G., & McGlynn, B. L. (2011). Hierarchical controls on runoff generation: Topographically driven hydrologic connectivity, geology, and vegetation. *Water Resources Research*, *47*(11), W11527. https://doi.org/10.1029/2011WR010666

Jencso, K. G., McGlynn, B. L., Gooseff, M. N., Wondzell, S. M., Bencala, K. E., & Marshall, L. A. (2009). Hydrologic connectivity between landscapes and streams: Transferring reach- and plot-scale understanding to the catchment scale. *Water Resources Research*, *45*(4), n/a–n/a). https://doi.org/10.1029/2008WR007225

Josefsson, W., & Landelius, T. (2000). Effect of clouds on UV irradiance: As estimated from cloud amount, cloud type, precipitation, global radiation and sunshine duration. *Journal of Geophysical Research: Atmospheres*, *105*(D4), 4927–4935. https://doi.org/10.1029/1999JD900255

Katsaros, K. B., McMurdie, L. A., Lind, R. J., & DeVault, J. E. (1985). Albedo of a water surface, spectral variation, effects of atmospheric transmittance, sun angle and wind speed. *Journal of Geophysical Research: Oceans*, *90*(C4), 7313–7321. https://doi.org/10.1029/JC090iC04p07313

Keys, T. A., Jones, C. N., Scott, D. T., & Chuquin, D. (2016). A cost-effective image processing approach for analyzing the ecohydrology of river corridors. *Limnology and Oceanography: Methods*, *14*(6), 359–369. https://doi.org/10.1002/lom3.10095

Kirkby, M. J. (1975). *Hydrograph modelling strategies*. Leeds, England: Department of Geography, University of Leeds.

Kubat, M., Holte, R. C., & Matwin, S. (1998). Machine learning for the detection of oil spills in satellite radar images. *Machine Learning*, *30*(2–3), 195–215. https://doi.org/10.1023/A:1007452223027

Kulasova, A., Beven, K. J., Blazkova, S. D., Rezacova, D., & Cajthaml, J. (2014). Comparison of saturated areas mapping methods in the Jizera Mountains, Czech Republic. *Journal of Hydrology and Hydromechanics*, *62*(2), 160–168. https://doi.org/10.2478/johh-2014-0002

Lang, R., Shao, G., Pijanowski, B. C., & Farnsworth, R. L. (2008). Optimizing unsupervised classifications of remotely sensed imagery with a data-assisted labeling approach. *Computers & Geosciences*, *34*(12), 1877–1885. https://doi.org/10.1016/j.cageo.2007.10.011

de Lima, J., P, o L. M., Abrantes, J., B, o R. C., , Silva, V. P., de Lima, M. I. P., & Montenegro, A. A. A. (2014). Mapping soil surface macropores using infrared thermography: An exploratory laboratory study. *The Scientific World Journal*, *2014*, e845460. https://doi.org/10.1155/2014/845460

de Lima, R. L. P., Abrantes, J. R. C. B., de Lima, J. L. M. P., & de Lima, M. I. P. (2015). Using thermal tracers to estimate flow velocities of shallow flows: Laboratory and field experiments. *Journal of Hydrology and Hydromechanics*, *63*(3), 255–262. https://doi.org/10.1515/johh-2015-0028

McDonnell, J. J. (2003). Where does water go when it rains? Moving beyond the variable source area concept of rainfall-runoff response. *Hydrological Processes*, *17*(9), 1869–1875. https://doi.org/10.1002/hyp.5132

McGlone, J. C., Mikhail, E. M., Bethel, J. S., & Sensing, A. S. for P. and R. (2004). *Manual of photogrammetry*. State College, PA: American Society for Photogrammetry and Remote Sensing.

Meerkerk, A. L., van Wesemael, B., & Bellin, N. (2009). Application of connectivity theory to model the impact of terrace failure on runoff in semi-arid catchments. *Hydrological Processes*, *23*(19), 2792–2803. https://doi.org/10.1002/hyp.7376

Mirus, B. B., Ebel, B. A., Heppner, C. S., & Loague, K. (2011). Assessing the detail needed to capture rainfall-runoff dynamics with physics-based hydrologic response simulation. *Water Resources Research*, *47*(3), W00H10. https://doi.org/10.1029/2010WR009906

Moody, J. A., & Martin, R. G. (2015). Measurements of the initiation of post-wildfire runoff during rainstorms using in situ overland flow detectors. *Earth Surface Processes and Landforms*, *40*(8), 1043–1056. https://doi.org/10.1002/esp.3704

Mundy, E., Gleeson, T., Roberts, M., Baraer, M., & McKenzie, J. M. (2017). Thermal imagery of groundwater seeps: Possibilities and limitations. *Groundwater*, *55*(2), 160–170. https://doi.org/10.1111/gwat.12451

Orlandini, S., Moretti, G., Corticelli, M. A., Santangelo, P. E., Capra, A., Rivola, R., & Albertson, J. D. (2012). Evaluation of flow direction methods against field observations of overland flow dispersion. *Water Resources Research*, *48*(10), W10523. https://doi.org/10.1029/2012WR012067

Parajka, J., Haas, P., Kirnbauer, R., Jansa, J., & Blöschl, G. (2012). Potential of time-lapse photography of snow for hydrological purposes at the small catchment scale. *Hydrological Processes*, *26*(22), 3327–3337. https://doi.org/10.1002/hyp.8389

Pawlowicz, R. (2003). Quantitative visualization of geophysical flows using low-cost oblique digital time-lapse imaging. *IEEE Journal of*

Oceanic Engineering, *28*(4), 699–710. https://doi.org/10.1109/JOE.2003.819310

Peñuela, A., Javaux, M., & Bielders, C. L. (2015). How do slope and surface roughness affect plot-scale overland flow connectivity? *Journal of Hydrology*, *528*, 192–205. https://doi.org/10.1016/j.jhydrol.2015.06.031

Pfister, L., McDonnell, J. J., Hissler, C., & Hoffmann, L. (2010). Ground-based thermal imagery as a simple, practical tool for mapping saturated area connectivity and dynamics. *Hydrological Processes*, *24*(21), 3123–3132. https://doi.org/10.1002/hyp.7840

Provost, F., & Kohavi, R. (1998). Guest editors' introduction: On applied research in machine learning. *Machine learning*, 30(2-3), 127–132. https://doi.org/10.1023/A:1007442505281

Reaney, S. M., Bracken, L. J., & Kirkby, M. J. (2014). The importance of surface controls on overland flow connectivity in semi-arid environments: Results from a numerical experimental approach. *Hydrological Processes*, *28*(4), 2116–2128. https://doi.org/10.1002/hyp.9769

Rogger, M., Pirkl, H., Viglione, A., Komma, J., Kohl, B., Kirnbauer, R., … Blöschl, G. (2012). Step changes in the flood frequency curve: Process controls. *Water Resources Research*, *48*(5), W05544. https://doi.org/10.1029/2011WR011187

Röper, T., Greskowiak, J., & Massmann, G. (2014). Detecting small groundwater discharge springs using handheld thermal infrared imagery. *Groundwater*, *52*(6), 936–942. https://doi.org/10.1111/gwat.12145

Rossi, M. J., & Ares, J. O. (2012). Depression storage and infiltration effects on overland flow depth-velocity-friction at desert conditions: Field plot results and model. *Hydrology and Earth System Sciences*, *16*(9), 3293–3307. https://doi.org/10.5194/hess-16-3293-2012

Salvatori, R., Plini, P., Giusto, M., Valt, M., Salzano, R., Montagnoli, M., … Sigismondi, D. (2011). Snow cover monitoring with images from digital camera systems. *Italian Journal of Remote Sensing*, 137–145. https://doi.org/10.5721/ItJRS201143211

Schulz, K., Seppelt, R., Zehe, E., Vogel, H. J., & Attinger, S. (2006). Importance of spatial structures in advancing hydrological sciences. *Water Resources Research*, *42*(3), n/a–n/a). https://doi.org/10.1029/2005WR004301

Shapiro, L. G., & Stockman, G. C. (2001). *Computer vision* (1st ed.). Upper Saddle River, NJ: Pearson.

Skøien, J. O., Blöschl, G., & Western, A. W. (2003). Characteristic space scales and timescales in hydrology. *Water Resources Research*, *39*(10), n/a–n/a). https://doi.org/10.1029/2002WR001736

Smits, P. C., Dellepiane, S. G., & Schowengerdt, R. A. (1999). Quality assessment of image classification algorithms for land-cover mapping: A review and a proposal for a cost-based approach. *International Journal of Remote Sensing*, *20*(8), 1461–1486. https://doi.org/10.1080/014311699212560

Solomons, A. G., Mikhailova, E. A., Post, C. J., & Sharp, J. L. (2015). LiDAR-based predictions of flow channels through riparian buffer zones. *Water Science*, *29*(2), 123–133. https://doi.org/10.1016/j.wsj.2015.11.001

Sproul, A. B. (2007). Derivation of the solar geometric relationships using vector analysis. *Renewable Energy*, *32*(7), 1187–1205. https://doi.org/10.1016/j.renene.2006.05.001

Tetzlaff, D., Carey, S. K., Laudon, H., & McGuire, K. (2010). Catchment processes and heterogeneity at multiple scales—Benchmarking observations, conceptualization and prediction. *Hydrological Processes*, *24*(16), 2203–2208. https://doi.org/10.1002/hyp.7784

Thenkabail, P. S. (2015). *Remotely sensed data characterization, classification, and accuracies*. Boca Raton, FL: CRC Press.

Tiner, R., Mulamoottil, G., Warner, B., & McBean, E. (1996). Practical considerations for wetland identification and boundary delineation. In *Wetlands environmental gradients, boundaries, and buffers* (pp. 113–137). Boca Raton, FL, USA: Lewis Publishers.

Vertessy, R. A., & Elsenbeer, H. (1999). Distributed modeling of storm flow generation in an Amazonian rain forest catchment: Effects of model parameterization. *Water Resources Research*, *35*(7), 2173–2187. https://doi.org/10.1029/1999WR900051

Weill, S., Altissimo, M., Cassiani, G., Deiana, R., Marani, M., & Putti, M. (2013). Saturated area dynamics and streamflow generation from coupled surface–subsurface simulations and field observations. *Advances in Water Resources*, *59*, 196–208. https://doi.org/10.1016/j.advwatres.2013.06.007

Western, A. W., Blöschl, G., & Grayson, R. B. (1998). How well do indicator variograms capture the spatial connectivity of soil moisture? *Hydrological Processes*, *12*(12), 1851–1868. https://doi.org/10.1002/(SICI)1099-1085(19981015)12:12<1851::AID-HYP670>3.0.CO;2-P

Western, A. W., Blöschl, G., & Grayson, R. B. (2001). Toward capturing hydrologically significant connectivity in spatial patterns. *Water Resources Research*, *37*(1), 83–97. https://doi.org/10.1029/2000WR900241

Weyman, D. R. (1973). Measurements of the downslope flow of water in a soil. *Journal of Hydrology*, *20*(3), 267–288. https://doi.org/10.1016/0022-1694(73)90065-6

Yang, J., & Chu, X. (2013). Quantification of the spatio-temporal variations in hydrologic connectivity of small-scale topographic surfaces under various rainfall conditions. *Journal of Hydrology*, *505*, 65–77. https://doi.org/10.1016/j.jhydrol.2013.09.013

Zha, Y., Gao, J., & Ni, S. (2003). Use of normalized difference built-up index in automatically mapping urban areas from TM imagery. *International Journal of Remote Sensing*, *24*(3), 583–594. https://doi.org/10.1080/01431160304987

How to cite this article: Silasari R, Parajka J, Ressl C, Strauss P, Blöschl G. Potential of time-lapse photography for identifying saturation area dynamics on agricultural hillslopes. *Hydrological Processes*. 2017;31:3610–3627. https://doi.org/10.1002/hyp.11272